

Message

---

**From:** LEE, LILY [/O=EXCHANGELABS/OU=EXCHANGE ADMINISTRATIVE GROUP (FYDIBOHF23SPDLT)/CN=RECIPIENTS/CN=D6085A744F9347E6836C54C0E85B97B2-LLEE06]  
**Sent:** 8/22/2016 6:25:50 PM  
**To:** Nguyen, Lyndsey [Nguyen.Lyndsey@epa.gov]; Rob Terry (Terry.Robert@epa.gov) [Terry.Robert@epa.gov]; Walker, Stuart [Walker.Stuart@epa.gov]  
**Subject:** For this afternoon's call  
**Attachments:** RMAC-0809-0012-0052 Fnl SUPRA NORM\_CD.pdf

Dear Lyndsey, Rob, and Stuart,

Thank you for agreeing to be on the call this afternoon. I wanted to ask if this could be one of the topics that we discuss. The attached pdf p.42 has a description of the Navy's approach. I've cut & pasted below, but the formatting got messed up. Thanks!

- Lily

Lily Lee  
Cleanup Project Manager  
Superfund Division  
U.S. Environmental Protection Agency, Region 9  
75 Hawthorne St. (SFD-8-3)  
San Francisco, CA 94105  
Tel: 415-947-4187, Fax: 415-947-3518  
[www.epa.gov/region9/superfund](http://www.epa.gov/region9/superfund)

---

**From:** LEE, LILY  
**Sent:** Thursday, August 04, 2016 3:38 PM  
**To:** Nguyen, Lyndsey <Nguyen.Lyndsey@epa.gov>  
**Subject:** Method used by Navy

## 5.0 STATISTICAL TESTS

As stated in Section 8.2.2.1 of MARSSIM (DoD et al. 2000), there is no need to conduct statistical tests when all sample results are less than the release criteria. Because all FSS sample results for the SUPRs were less than the release criteria, no statistical tests were included in the SUPRs. However, this section describes the process by which the sampling density was determined such that statistical testing, if necessary, would be sufficient.

### 5.1 DECISION ERRORS

There are two types of decision errors that were used when performing the statistical tests outlined in MARSSIM (DoD et al. 2000). The first type of decision error, called a Type I error, occurs when the null hypothesis is rejected when it is actually true. This type of error is sometimes called a "false positive." The probability of a Type I error is denoted by an  $\alpha$ . The Type I error is often referred to as the significance level or size of the test.

The second type of decision error, called a Type II error, occurs when the null hypothesis is not rejected when it is actually false. This type of error is sometimes called a "false negative." The probability of a Type II error is denoted by a  $\beta$ . The power of a statistical test is defined as the probability of rejecting the null hypothesis when it is actually false. It is numerically equal to  $1 - \beta$ , where  $\beta$  is the Type II error rate.

Each survey was designed to limit Type I and Type II errors to a maximum probability of 5 percent. It was important to minimize the chances of concluding that a survey unit met the release criteria (reject the null hypothesis) when it actually exceeded the limits (Type I error), and concluding that

a survey unit exceeded the release criteria (accept the null hypothesis) when it actually met the limits (Type II error).

## 5.2 WILCOXON RANK-SUM TEST

The Wilcoxon Rank-Sum (WRS) test is designed to test a hypothesis about the location of a population distribution. It is most often used to test the hypothesis about a population median and often involves the use of matched pairs. For example, reference area and survey unit data were tested for a median difference of zero. This test is also a nonparametric test that may be used when it is only necessary, or possible, to know if observed differences between two conditions were significant. The WRS test is structured to denote a change in magnitude, as opposed to any attempt at the quantitative measurement. Per Section 2.5.1.2 of MARSSIM (DoD et al. 2000), the WRS test is the recommended statistical test for comparison of survey unit radionuclide concentrations with background radionuclide concentrations. Although the WRS test was not conducted for the SUPRs because the results for all survey units were less than the release criteria, the number of sampling points was calculated to ensure that sufficient sample results from the survey unit areas were available for statistical comparison had all FSS results not all been lower than each release criterion.

### 5.2.1 Determining the Numbers of Data Points for the WRS Test

Since the ROCs were present in background,  $N$  is calculated in the manner specified for the WRS test using Equation 5-1:

**Equation 5-1**

$$\left( \frac{Z_{1-\alpha} + Z_{1-\beta}}{P_r} \right)^2 (1.2) = N$$

Where:

$Z_{1-\alpha}$  = 1.645 Type I decision error level

$Z_{1-\beta}$  = 1.645 Type II decision error level

$P_r$  = 0.997658 random measurement probability

(from Table 5.1 in MARSSIM [NUREG-1575; DoD et al. 2000])

(1.2) = 20% increase in number of samples over the minimum

The variable used to calculate  $N$  not already specified in the Basewide Radiological Management Plan (TtEC 2012a) was the random measurement probability  $P_r$ .  $P_r$  in Equation 5-1 above is based on the relative shift. The relative shift is equal to  $\Delta/\sigma$ , where  $\Delta$  was equal to the release criterion minus the lower boundary of the gray region (LBGR), and  $\sigma$  was an estimate of the standard deviation of the measured values in a survey unit. The LBGR was the net median concentration of the contaminant in the survey unit. When this value was unknown, half the value of the Derived Concentration Guideline Limit corresponding to a total effective dose equivalent of 25 mrem/y to

a member of the critical group was used as the LBGR, as suggested in MARSSIM (DoD et al. 2000). Likewise,  $\sigma$  was assigned the value of the standard deviation of the measurement values in the reference area.

Using Equation 5-2 below, the value for  $\Delta$  was derived by converting the release criterion from pCi/g to counts per second (cps). To perform this conversion, an arbitrary concentration of  $^{226}\text{Ra}$  is divided by the associated exposure rate produced by this concentration of  $^{226}\text{Ra}$ . The resulting number was then divided by the average net cps per microrentgen per hour ( $\mu\text{R/h}$ ) for the detectors on the towed array. Once this number was derived, the release criterion of 1.0 pCi/g for  $^{226}\text{Ra}$  was divided by this number.

**Equation 5-2**

$AC \cdot ER \cdot CR$

$RC \cdot \text{Cin} \cdot \text{cpm} \cdot RC$

$/ * 1 /$

=

Where:

$RC$  = release criterion (pCi/g)

$AC$  = arbitrary concentration of  $^{226}\text{Ra}$  (pCi/g)

$ER$  = exposure rate for 1 pCi/g  $^{226}\text{Ra}$  calculated by Microshield®

$CR$  = counts per second per  $\mu\text{R/h}$  for the detector

Where:

$RC = 1.0$

$AC = 1$

$ER = 0.6984$

$CR = 436$

The values used for these parameters to determine a  $P_r$  were 304 cps for the release criterion.

Taking half of this value, 152 cps for  $\Delta$ , and using 30 cps for  $\sigma$ ,  $N$  was calculated as nine samples when using Equation 5-1. To obtain reasonable assurance that any small areas of elevated residual radioactivity were not missed during the survey, the total number of samples was increased by a factor of two, for a total of 18 samples in each survey unit.

RMAC-

Lily Lee

Cleanup Project Manager

Superfund Division

U.S. Environmental Protection Agency, Region 9

75 Hawthorne St. (SFD-8-3)

San Francisco, CA 94105

Tel: 415-947-4187, Fax: 415-947-3518

[www.epa.gov/region9/superfund](http://www.epa.gov/region9/superfund)